

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: May 20, 2005, 19:25:58 ; Search time 165 Seconds  
(without alignments)  
1924.426 Million cell updates/sec

Title: ALLEN-724676-70493

Perfect score: 4338  
Sequence: 1 MDVVDPDIFNRDPDRDHTDL.....VETRPVDPDPAPSNIYIQE 821Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

## Database :

A\_Geneseq\_16Dec04:\*

1: geneseqp19808:\*

2: geneseqp19908:\*

3: geneseqp20008:\*

4: geneseqp20018:\*

5: geneseqp20028:\*

6: geneseqp20038:\*

7: geneseqp20038:\*

8: geneseqp20048:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4288	98.8	847	8	ADR39751 Human kin
2	4196	96.7	833	6	AAO30330 Human MAP
3	4196	96.7	833	8	ADM72219 Human TAS
4	4196	96.7	833	8	ADP24693 PRO polyp
5	4057	93.5	772	4	ABG01438 Novel hum
6	2421	55.8	495	7	ADB37566 Neural th
7	2421	55.8	495	7	ADJ59822 Human hea
8	1930	44.5	842	2	AAVS5956 Human STE
9	1930	44.5	846	6	AAO30335 Human MAP
10	1930	44.5	846	6	ADP23570 PRO polyp
11	1917.5	44.2	873	6	ABG74448 Rat Germ
12	1904.5	43.9	873	4	AAW78346 Human pro
13	1898	43.8	894	2	AAVS5935 Human KHS
14	1898	43.8	894	6	AAO30333 Human pro
15	1898	43.8	894	8	ADP24927 PRO polyp
16	1896	43.7	894	4	AAW78345 Human pro
17	1896	43.7	930	4	AAW79329 Human pro
18	1896	43.7	930	4	AAW79330 Human pro
19	1877	42.3	884	6	AAO30334 Human MAP
20	1840.5	42.4	855	4	ABG20230 Novel hum
21	1697	39.1	820	6	AAO30332 Human MAP
22	1694	39.1	819	6	AAO30331 Human MAP
23	1694	39.1	819	8	ADL12640 Human ste
24	1518	35.0	1218	4	ABB62736 Drosophila
25	1444	33.3	698	8	ADN61489 Human KPP

26	1394.5	32.1	644	8	ABM63565 Human dia
27	717.5	16.5	979	4	ABB66102 Drosophila
28	675	15.6	1109	2	AAVS5953 Nematode
29	668.5	15.4	911	2	AAVS5957 Mouse STE
30	663.5	15.3	1286	8	ADK71845 Human kin
31	663.5	15.3	1295	5	AAO18508 Human ins
32	663.5	15.3	1295	7	ADG36898 Human GCK
33	663.5	15.3	1295	8	ADG36898 Human GCK
34	663.5	15.3	1303	4	AAW79153 Human pro
35	663.5	15.3	1303	5	AAO18507 Human ins
36	663.5	15.3	1303	7	ADG36900 Human GCK
37	661.5	15.2	1135	8	ADR39750 Human kin
38	661.5	15.2	1303	4	AAE10612 Human nov
39	660.5	15.2	1312	7	ADBE34150 Human mls
40	651	15.0	449	6	ABP96448 STR4/Met-
41	651	15.0	462	8	AD140883 Human kin
42	651	15.0	462	8	ADL70266 LBRFL305 p
43	651	15.0	467	8	ADL70268 Human Mbc
44	651	15.0	487	8	ADR89878 Human MBC
45	650	15.0	426	2	AAW31603 Human pro

## ALIGNMENTS

RESULT 1	
ADR39751	
ID	ADR39751 standard; protein; 847 AA.
XX	
AC	ADR39751;
XX	
DT	18-NOV-2004 (first entry)
XX	
DE	Human kinase and phosphatase KPP-24 protein SEQ ID NO:24.
XX	
XX	human; kinase and phosphatase protein; KPP; enzyme; cytosolic;
KW	antiarteriosclerotic; anticonvulsant; nootropic; neuroprotective;
KW	cardioprotective; anti-HIV; antiallergic; antiinflammatory;
KW	thymimetic; gene therapy; cell proliferative disorder; cancer;
KW	atherosclerosis; neurological disorder; epilepsy; Huntington's disease;
KW	stroke; immune disorder; inflammatory disorder; AIDS; allergy;
KW	developmental disorder; Hypothyroidism; Cushing's syndrome; infection;
KW	KPP-24.
XX	
OS	Homo sapiens.
XX	
PN	WO2004074453-A2.
XX	
PD	02-SEP-2004.
XX	
PF	20-FEB-2004; 2004WO-US005092.
XX	
PR	20-FEB-2003; 2003US-0449059P.
XX	
PR	19-MAR-2003; 2003US-0456932P.
XX	
PR	28-MAR-2003; 2003US-0458844P.
XX	
PR	09-APR-2003; 2003US-0461678P.
XX	
PR	17-APR-2003; 2003US-0463937P.
XX	
PA	(INCY-) INCYTE CORP.
XX	
PI	Ramkumar J, Marquis JP, Swarnakar A, Chawla NK, Tran UK;
XX	Becha SD, Lee SY, Hafalia AJA, Richardson TW, Khare R, Jiang X;
PI	Jackson AA, Yang J, Gorvad AE;
XX	
DR	N-PDB; ADR39797.
XX	
DR	WPI; 2004-635568/61.
XX	
PT	New human kinases and phosphatases (KPP) for diagnosing, treating and
XX	preventing diseases or conditions associated with aberrant KPP expression
PT	e.g. cancer, acquired immunodeficiency syndrome, epilepsy, or infections.
XX	
PS	Claim 1, SEQ ID NO 24; 29pp; English.
XX	

Result No.	Score	Query Match	Length	DB	ID	Description
1	1694	39.1	819	2	A537114	protein kinase (EC 2.7.1.1)
2	1291	29.8	829	2	T293772	hypothetical protein
3	710.5	16.4	1102	2	JC6316	probable protein kinase
4	693.5	16.0	836	2	B96716	probable serine/threonine kinase
5	645	14.9	426	2	S71866	Ste20-like protein
6	639	14.7	1075	2	T27623	hypothetical protein
7	638.5	14.6	653	2	T34356	hypothetical protein
8	623.5	14.4	1080	2	T27622	hypothetical protein
9	624.5	14.4	1233	2	T30989	hypothetical protein
10	604.5	13.9	680	2	C96572	serine/threonine kinase
11	580.5	13.4	471	2	T39232	protein P12M16.4
12	575	13.3	1233	2	T14157	probable serine/threonine kinase
13	574	13.0	1206	2	T34021	serine/threonine kinase
14	562	13.0	1228	2	T18897	protein kinase SK2
15	556	12.8	1231	2	T18532	hypothetical protein
16	551	12.7	652	2	T39722	serine/threonine protein kinase
17	544.5	12.6	312	2	T38525	serine/threonine kinase
18	534.5	12.3	490	2	S47946	serine/threonine kinase
19	529.5	12.2	1001	2	T11765	serine/threonine kinase
20	515	11.9	982	2	T18576	serine/threonine kinase
21	512	11.8	544	2	S40482	serine/threonine kinase
22	512	11.8	545	2	G01773	p21-activated protein kinase
23	508	11.7	1080	2	S48944	hypothetical protein
24	505.5	11.7	544	2	I49376	p21-activated protein kinase
25	504.5	11.6	525	2	S58682	protein kinase, p21-activated
26	500.5	11.5	545	2	A57597	beta-p21-activated protein kinase
27	498.5	11.5	658	2	T39600	serine/threonine kinase
28	478	11.0	939	2	S28394	probable serine/threonine kinase
29	473.5	10.9	658	2	S60170	protein kinase Pak1

30	473.5	10.9	1230	2	T18256	probable serine/thr
31	473.5	10.9	1230	2	T18259	serine/threonine p
32	468	10.8	608	2	G96575	probable MEK kinase
33	465.5	10.7	1062	2	S46316	protein kinase CDC
34	459	10.6	113	2	I38216	protein-serine/thr
35	457	10.5	622	2	T15457	hypothetical prote
36	456	10.5	589	2	T18086	serine/threonine-p
37	447.5	10.3	1014	2	T31109	myosin III - Actin
38	444	10.2	1135	1	A29813	13k minac protein
39	444	10.2	1501	1	B29813	17k minac protein
40	443	10.2	883	2	A96662	hypothetical prote
41	442	10.2	378	2	T26684	hypothetical prote
42	442	10.2	553	2	T01479	hypothetical prote
43	440.5	10.2	655	2	S51884	probable protein k
44	439	10.1	842	2	S60402	protein kinase CLA
45	433.5	10.0	487	2	A71407	probable Ste20-1lk

## ALIGNMENTS

## RESULT 1

protein kinase (EC 2.7.1.37) BL44 - human  
N.Alternate names: GC kinase  
C.Species: Homo sapiens (man)  
C.Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 09-Jul-2004  
C.Accession: A53714  
R.Katz, P.; Whalen, G.; Kehrl, J.H.  
J. Biol. Chem. 265, 16802-16809, 1994  
A.Title: Differential expression of a novel protein kinase in human B lymphocytes. Pref.  
A.Reference number: A53714; MUID:94266900; PMID:7515885  
A.Accession: A53714  
A.Status: Preliminary  
A.Molecule type: mRNA  
A.Residues: 1-819 <KAT>  
C.Cross-references: UNIPROT:Q12851; GB:U07349; NID:G531819; PIDN:AAA20968.1; PID:G53182060  
C.Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C.Keywords: ATP, phosphotransferase  
F.13-2/7/Domain: protein kinase ATP-binding motif  
F.21-29/Region: protein kinase ATP-binding motif

Query Match	39.1%;	Score 1694;	DB 2;	Length 819;
Best Local Similarity	41.8%;	Pred. No. 1.1e-61;		
Matches 362;	Conservative 142;	Mismatches 250;	Indels 112;	Gaps 18;

```

Qy 7 DIFR2DRDHDHLLORLGGGVEVFAKPKVSDVALKVVNNEPDDVSTLKEKILL 66
Db 5 DVSLQDDPRDRRELLQVRVAGTGDVYRARDLVISELAAVKVLDPGDSDLSLOEITLL 64
Qy 67 KTCCHNANIVAHGSGYTLWOKLWICMEFCCGASLODIDYQVGTSLSELISYVCREVLQSLA 126
Db 65 RECHPRVAVAYIGSYLNDRLMTIMECCGGSGLOEITHAGPBLERQIAYCRRRLNGLH 124
Qy 127 YLH5OKTIHRDIKANILLINDAGEVRLADFGISAQICATLAPRLSFGTPYMAAERYAAV 186
Db 125 HLHSQGIKIHDIKANILLITLQGVKLADGVSGSELTA5AKRSPFGTPYMAAERYAAV 184
Qy 187 ALKGSVNELEDISIGITATELAELOPRLPDVPHVRLFLMTSGVCPRLKEKGKSSAA 246
Db 185 ERKGSVNELEDVMAIGITALELOPRLPHLHPRLMTMSXSSFOPRKLDNRKTWTON 244
Qy 247 FHNFIKTYLTKSPKPKPSATKMLSHQV50PQGNRGLILDLDKLKNPKGK-PSIGIED 305
Db 245 FHNHLKATLKNPKRPTAKLQHPRTTQ-LPRALLTQLLDKASPHIGTSPREDCOL 303
Qy 306 BEBELPALPRIRISTHRSSLSIGIPDADCCRHMEPRKINGMETRPANTARLQRP--- 361
Db 304 EYDYM---PDTIHSRQGHGPAERTSEI0FHQVKEGAPAREKETDP-----LNEPWEEE 354
Qy 362 -----RDLRSSPKQJSESSDDYDVVDITPAADTP-----PPLP 398
Db 355 WTLIGKEELSGLLQ5VQALEERSLITRSASFGQEDSP---DDTWGITKRAPLPGPLP 411

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2005, 19:38:56 ; Search time 142 Seconds  
(without alignments)  
1934.019 Million cell updates/sec

Title: ALLEN-724676-70493

Perfect score: 4338  
Sequence: 1 MDVVDPDIFNRDPDRHYDL.....VETRPVDDPTAPSNLYIOE 821

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1434725 seqs, 334507595 residues

Total number of hits satisfying chosen parameters: 1434725

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

- Database : Published Applications AA:\*
- 1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/1/pubppaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/1/pubppaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/1/pubppaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/1/pubppaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/1/pubppaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/1/pubppaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/1/pubppaa/US10D\_PUBCOMB.pep.\*
  - 17: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/1/pubppaa/US11\_NEW\_PUB.pep.\*
  - 19: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep.\*
  - 20: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4303	99.2	836	15	US-10-425-114-39187 Sequence 39187, A
2	4196	96.7	833	14	US-10-303-683-17 Sequence 17, Appl
3	2421	55.8	495	14	US-10-198-070-48 Sequence 48, Appl
4	2421	55.8	495	16	US-10-408-765A-1628 Sequence 1628, Ap
5	1930	44.5	842	10	US-09-291-417-91 Sequence 91, Appl
6	1930	44.5	846	14	US-10-303-683-22 Sequence 22, Appl
7	1930	44.5	848	15	US-10-425-114-39246 Sequence 39246, A
8	1917.5	44.2	873	9	US-09-843-245-3 Sequence 3, Appl1
9	1898	43.8	894	10	US-09-291-417-18 Sequence 18, Appl
10	1898	43.8	894	14	US-10-303-683-20 Sequence 20, Appl
11	1877	43.3	884	14	US-10-303-683-21 Sequence 21, Appl
12	1697	39.1	820	14	US-10-303-683-19 Sequence 19, Appl
13	1694	39.1	819	14	US-10-303-683-18 Sequence 18, Appl

14	1014	23.4	322	16	US-10-664-421-78	Sequence 78, Appl
15	675	15.6	1109	10	US-09-291-417-88	Sequence 88, Appl
16	668.5	15.4	911	10	US-09-291-417-92	Sequence 92, Appl
17	663.5	15.3	1295	10	US-09-789-390-30	Sequence 30, Appl
18	663.5	15.3	1295	10	US-09-789-390-32	Sequence 32, Appl
19	663.5	15.3	1295	10	US-09-789-390-34	Sequence 34, Appl
20	663.5	15.3	1295	10	US-09-789-390-37	Sequence 37, Appl
21	663.5	15.3	1295	10	US-09-789-390-39	Sequence 39, Appl
22	663.5	15.3	1303	10	US-09-789-390-35	Sequence 35, Appl
23	663.5	15.3	1303	10	US-09-789-390-38	Sequence 38, Appl
24	661.5	15.2	1303	10	US-09-789-390-9	Sequence 9, Appl
25	660.5	15.2	1312	14	US-10-029-115-2	Sequence 2, Appl1
26	660.5	15.2	1312	17	US-10-493-164-1	Sequence 1, Appl1
27	650	15.0	1332	10	US-10-498-698-9	Sequence 9, Appl1
28	649	15.0	1332	10	US-09-789-390-7	Sequence 7, Appl1
29	648.5	14.9	1268	15	US-09-789-390-13	Sequence 13, Appl
30	647	14.9	1268	15	US-10-353-690-122	Sequence 122, App
31	647	14.9	1360	9	US-09-871-916-2	Sequence 2, Appl1
32	647	14.9	1360	14	US-10-355-975-14	Sequence 14, Appl
33	646	14.9	1212	14	US-10-247-671-157	Sequence 157, App
34	646	14.9	1212	15	US-10-168-582-9	Sequence 9, Appl1
35	645	14.9	426	9	US-09-906-397-4	Sequence 4, Appl1
36	645	14.9	426	9	US-09-862-027-52	Sequence 52, Appl
37	645	14.9	426	17	US-10-772-636-42	Sequence 42, Appl
38	645	14.9	426	17	US-10-989-228-52	Sequence 52, Appl
39	643.5	14.8	1165	14	US-10-021-660-126	Sequence 126, App
40	643.5	14.8	1165	15	US-10-211-462-89	Sequence 89, Appl
41	643	14.8	426	9	US-09-862-027-54	Sequence 54, Appl
42	643	14.8	426	10	US-09-291-417-84	Sequence 84, Appl
43	643	14.8	426	17	US-10-989-228-54	Sequence 84, Appl
44	642	14.8	487	9	US-09-810-808-8	Sequence 8, Appl1
45	640.5	14.8	1273	10	US-09-789-390-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1  
US-10-425-114-39187  
Sequence 39187, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E  
APPLICANT: Tabaska, Jack E  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21 (5313) B  
CURRENT APPLICATION NUMBER: US/10/425, 114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 39187  
LENGTH: 836  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB4676-011-A2\_FLI\_PEP  
US-10-425-114-39187

Query Match 99.2%; Score 4303; DB 15; Length 836;  
Best Local Similarity 99.4%; Pred. No. 1,4e-261;  
Matches 816; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 MDVVDPDIFNRDPDRHYDLQRLGGTYGVKFKARDKVGDIVALRMVXKPEPDDVSTLQ 60  
Db 16 MDVVDPDIFNRDPDRHYDLQRLGGTYGVKFKARDKVGDIVALRMVXKPEPDDVSTLQ 75  
QY 61 KRLILIKTRHANIYAVHGSYIMLQKIMCFPCGAGSLQDIYQVTSLSRLQISTYCKE 120  
Db 76 KRLILIKTRHANIYAVHGSYIMLQKIMCFPCGAGSLQDIYQVTSLSRLQISTYCKE 135

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2005, 19:31:59 ; Search time 25 Seconds  
(without alignments)  
2451.475 Million cell updates/sec

Title: ALLEN-724676-70493

Perfect score: 4338  
Sequence: 1 MDVVDPIENRDPDRHYDL.....VETRPVDDPTAPSNLYIQE 821

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiletest.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4196	96.7	847	4	US-09-949-016-11002 Sequence 11002, A
2	1930	44.5	842	4	US-09-688-188B-91 Sequence 91, Appl
3	1930	44.5	842	4	US-09-291-417D-91 Sequence 91, Appl
4	1898	43.8	894	4	US-09-688-188B-18 Sequence 18, Appl
5	1898	43.8	894	4	US-09-291-417D-18 Sequence 18, Appl
6	1697	39.1	827	4	US-09-949-016-7807 Sequence 7807, Ap
7	1694	39.1	819	4	US-09-976-594-369 Sequence 369, App
8	977.5	22.5	276	2	US-08-852-743-7 Sequence 7, Appl
9	977.5	22.5	276	3	US-09-185-370-7 Sequence 7, Appl
10	681	15.7	463	4	US-09-270-767-43524 Sequence 43524, A
11	675	15.6	1109	4	US-09-688-188B-88 Sequence 88, Appl
12	675	15.6	1109	4	US-09-291-417D-88 Sequence 88, Appl
13	668.5	15.4	911	4	US-09-688-188B-92 Sequence 92, Appl
14	668.5	15.4	911	4	US-09-291-417D-92 Sequence 92, Appl
15	668.5	15.4	911	4	US-09-688-188B-154 Sequence 154, App
16	668.5	15.4	966	4	US-09-291-417D-154 Sequence 154, App
17	647	14.9	1269	4	US-09-645-456A-15 Sequence 15, Appl
18	647	14.9	1269	4	US-09-425-324A-15 Sequence 15, Appl
19	647	14.9	1269	4	US-09-645-791-15 Sequence 15, Appl
20	647	14.9	1269	4	US-09-645-456A-12 Sequence 12, Appl
21	647	14.9	1277	4	US-09-425-324A-12 Sequence 12, Appl
22	647	14.9	1277	4	US-09-645-791-12 Sequence 12, Appl
23	647	14.9	1298	4	US-09-645-456A-14 Sequence 14, Appl
24	647	14.9	1298	4	US-09-425-324A-14 Sequence 14, Appl
25	647	14.9	1298	4	US-09-645-791-14 Sequence 14, Appl
26	647	14.9	1306	4	US-09-645-456A-10 Sequence 10, Appl
27	647	14.9	1306	4	US-09-425-324A-10 Sequence 10, Appl

28	647	14.9	1306	4	US-09-645-791-10 Sequence 10, Appl
29	647	14.9	1324	4	US-09-645-456A-13 Sequence 13, Appl
30	647	14.9	1324	4	US-09-425-324A-13 Sequence 13, Appl
31	647	14.9	1324	4	US-09-645-791-13 Sequence 13, Appl
32	647	14.9	1332	4	US-09-645-456A-9 Sequence 9, Appl
33	647	14.9	1332	4	US-09-425-324A-9 Sequence 9, Appl
34	647	14.9	1332	4	US-09-645-791-9 Sequence 9, Appl
35	647	14.9	1353	4	US-09-645-456A-11 Sequence 11, Appl
36	647	14.9	1353	4	US-09-425-324A-11 Sequence 11, Appl
37	647	14.9	1353	4	US-09-645-791-11 Sequence 11, Appl
38	647	14.9	1360	3	US-09-393-569-2 Sequence 2, Appl
39	647	14.9	1360	4	US-09-579-664B-14 Sequence 14, Appl
40	647	14.9	1360	4	US-09-645-456A-34 Sequence 34, Appl
41	647	14.9	1360	4	US-09-425-324A-34 Sequence 34, Appl
42	647	14.9	1360	4	US-09-645-791-34 Sequence 34, Appl
43	647	14.9	1360	4	US-10-355-975A-14 Sequence 14, Appl
44	645	14.9	426	2	US-08-852-743-2 Sequence 2, Appl
45	645	14.9	426	2	US-09-211-930-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1		US-09-949-016-11002		Application US/09949016	
Sequence 11002, Appl		Patent No. 6812339			
GENERAL INFORMATION:					
APPLICANT: VENTER, J. Craig et al.					
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED					
FILE REFERENCE: C1001307					
CURRENT APPLICATION NUMBER: US/09/949, 016					
PRIOR FILING DATE: 2000-04-14					
PRIOR APPLICATION NUMBER: 60/241, 755					
PRIOR FILING DATE: 2000-10-20					
PRIOR APPLICATION NUMBER: 60/237, 768					
PRIOR FILING DATE: 2000-10-03					
PRIOR APPLICATION NUMBER: 60/231, 498					
PRIOR FILING DATE: 2000-09-08					
NUMBER OF SEQ ID NOS: 207012					
SOFTWARE: FASTSEQ for Windows Version 4.0					
SEQ ID NO 11002					
LENGTH: 847					
TYPE: PRT					
ORGANISM: Human					
US-09-949-016-11002					
Query Match		96.7%; Score 4196; DB 4; Length 847;			
Best Local Similarity		99.5%; Pred. No. 0;			
Matches 795; Conservative		0; Mismatches 4; Indels 0; Gaps 0;			
QY	1	MDVVDPIENRDPDRHYDLQRLGGTGYGVFARQVSGDLVAKVKNKPPDDVSTLQ	60		
DB	15	MDVVDPIENRDPDRHYDLQRLGGTGYGVFARQVSGDLVAKVKNKPPDDVSTLQ	74		
QY	61	KEILLIKTCHNANIVVHGSYVWLQKLMICMECGASLQDIYQVGSLSLQISVCRE	120		
DB	15	MDVVDPIENRDPDRHYDLQRLGGTGYGVFARQVSGDLVAKVKNKPPDDVSTLQ	74		
QY	75	KEILLIKTCHNANIVVHGSYVWLQKLMICMECGASLQDIYQVGSLSLQISVCRE	134		
QY	121	VLOGLAIVHSQKIHNDIKANILINDAGVRLADFGISAQIGANTLARRLSFGTPYMA	180		
DB	135	VLOGLAIVHSQKIHNDIKANILINDAGVRLADFGISAQIGANTLARRLSFGTPYMA	194		
QY	181	PEVAALVAKGQVNECDIWSLGTALBLAELQPLPDVHPLVLPMTKSGYOPPRILEK	240		
DB	195	PEVAALVAKGQVNECDIWSLGTALBLAELQPLPDVHPLVLPMTKSGYOPPRILEK	254		
QY	241	GKSAFAHNHFKITLTKSPKRSATKMLSHQVSPGALNRGLTLDLKLPKGPSPST	300		
DB	255	GKSAFAHNHFKITLTKSPKRSATKMLSHQVSPGALNRGLTLDLKLPKGPSPST	314		
QY	301	GDIEDEPELPPAIPRIRISTRSSSLGIPDADCCRRHMEFKLRGMETRPANTARLQ	360		

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: May 20, 2005, 19:26:34 ; Search time 177 Seconds  
(without alignments)  
2375.237 Million cell updates/sec

Title: ALLEN-724676-70493

Perfect score: 4338  
Sequence: 1 MDVVDPIFNRPDRHYDL.....VETRPVDDPTAPSNXTIQE 821

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: UniProt\_03: \*  
1: uniProt\_sprot: \*  
2: uniProt\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4196	96.7	833	1	MAK1_HUMAN
2	3707	85.5	827	1	MAK1_MOUSE
3	1947	44.9	878	2	OZU2
4	1930	44.5	846	1	MAK5_HUMAN
5	1905.5	43.9	847	1	MAK5_MOUSE
6	1898	43.8	862	1	MAK3_HUMAN
7	1882.5	43.4	823	2	O68F04
8	1694	39.1	819	1	MAK2_HUMAN
9	1688	38.9	812	2	O68V03
10	1688	38.5	821	1	MAK2_MOUSE
11	1652.5	38.1	881	1	MAK3_MOUSE
12	1633.5	37.7	947	2	O6ML18
13	1632.5	37.6	947	2	O6SV11
14	1518	35.0	1218	2	O9V8R6
15	1444	33.3	675	2	O7Z266
16	1404.5	32.4	632	2	O7Z299
17	1291	29.8	829	2	OQ3290
18	1005.5	23.2	318	2	O7QGN6
19	759	17.5	406	2	O6PEQ2
20	747	17.2	1042	2	O7KVB9
21	721.5	16.6	737	2	O68IX1
22	721	16.6	1200	2	O7KVB9
23	721	16.6	1504	2	O9W002
24	720	16.6	1102	2	O7KVB8
25	710.5	16.4	1102	2	O7M3J7
26	710.5	16.4	1534	1	O7PIC2
27	709	16.3	916	1	TNIK_MOUSE
28	704.5	16.2	855	2	O8LKT7
29	699.5	16.1	1120	2	O9LQW1
30	699	16.1	829	2	O871H9

32	699	16.1	842	2	O9FNU3
33	693.5	16.0	836	2	O245Z7
34	690	15.9	461	2	O61125
35	682	15.7	825	2	O8SAB1
36	681.5	15.7	825	2	O9ARL7
37	677.5	15.6	1082	2	O8TBM3
38	672	15.5	1087	2	O9XYC3
39	672	15.5	1096	2	O95Z17
40	668.5	15.4	966	1	STKA_MOUSE
41	666.5	15.4	1072	2	O814B5
42	661.5	15.2	1332	1	MAK6_HUMAN
43	660	15.2	809	2	O6VTC1
44	658	15.2	688	2	O6Y2W8
45	654.5	15.1	420	2	O7SYN3

## ALIGNMENTS

RESULT 1  
MAK1\_HUMAN STANDARD; PRT; 833 AA.  
AC Q92918  
BT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DE 85-JUN-2004 (Rel. 44, Last annotation update)  
DE Mitogen-activated protein kinase kinase kinase 1 (EC 2.7.1.37)  
DE (MAPK/ERK kinase kinase kinase 1) (MEK kinase kinase 1)  
DE (Hematopoietic progenitor kinase)  
GN Name=MAPK1; Synonyms=HPK1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCIT\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND INTERACTION WITH  
RP MAP3K1.  
RC TISSUE=Fetal liver;  
RX MEDLINE=96421968; PubMed=8824585;  
RZ Hu M.C.-T., Qiu W.R., Wang X., Meyer C.F., Tan T.-H.,  
RT "Human HPK1, a novel human hematopoietic progenitor kinase that  
RL Genes Dev. 10:2251-2264(1996).  
CC -1- FUNCTION: May play a role in the response to environmental stress.  
CC -1- Appears to act upstream of the JUN N-terminal pathway. May play a  
CC role in hematopoietic lineage decisions and growth regulation.  
CC -1- role in hematopoietic lineage decisions and growth regulation.  
CC -1- COFACTOR: Magnesium.  
CC -1- SUBUNIT: Interacts with MAP3K1.  
CC -1- TISSUE SPECIFICITY: Expressed primarily in hematopoietic organs,  
CC including bone marrow, spleen and thymus. Also expressed at very  
CC low levels in lung, kidney, mammary glands and small intestine.  
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family. STE20  
CC subfamily.  
CC -1- SIMILARITY: Contains 1 CNH domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
CC  
CC EMBL; U66464; AAB97983.1; -  
CC HSSP; Q13153; 1F3M.  
CC InAcct; Q92918; -  
CC Genew; HGNC:6863; MAPK1.  
CC MIM; 601983; -  
CC GO; GO:0005524; F:ATP binding; IDA.  
CC GO; GO:0004674; F:protein serine/threonine kinase activity; IDA.  
CC GO; GO:0007257; P:activation of JNK; TAS.  
CC GO; GO:0006468; P:protein amino acid phosphorylation; IDA.